

WEBSITE EMAIL ANTI-HARVESTING TECHNIQUES

Eddie Sparpaglione

Sussex County

What is Email Harvesting?

E-mail Harvesting is the process of obtaining lists of e-mail addresses using various methods for use in bulk e-mail or other purposes usually grouped as spam.

Anti-Harvesting Solutions

- List-poisoning
- Address Munging
- Client-Side Scripting (JavaScript)
- CSS Manipulation
- Images
- Contact Form

List Poisoning

Definition: poisoning a mailing list with invalid email addresses

□ Advantages

- ▣ Gives the spammer a list of fake email addresses which is useless

□ Disadvantages

- ▣ Causes problems for the mail server responsible for that address
- ▣ Can be avoided by spammers if they are found
- ▣ *Only practical for industries using mailing lists, not individual addresses*

Address Munging

Definition: disguising an e-mail address to prevent it from being collected

□ Example

□ No-one@example.com = no-one at example dot com

□ Advantages

- Deters simple bots
- Able to be seen directly by the user

□ Disadvantages

- Makes it more difficult for users to send email to each other
- Makes the user less likely to email because it adds another step
- Many bots can still decipher this technique

Client-Side Scripting

Definition: represent e-mail addresses as ASCII code or JavaScript-coded text

□ Advantages

- Spam-bots cannot find the address in the source code

□ Disadvantages

- JavaScript must be enabled
- Can cause problems for people with screen readers

Example:

Using JavaScript to represent user@example.com:

```
<script type="text/javascript">
var name = 'user';
var at = '@';
var domain = 'example.com';
document.write(name + at + domain);
</script>
```

CSS Manipulation

Definition: Listing the e-mail addresses in a scrambled format and having CSS decode them

□ Advantages

- ▣ Easily reusable
- ▣ Can create your own pattern therefore harder to harvest
- ▣ Cannot harvest the address from source code

□ Disadvantages

- ▣ Cannot copy and paste the address
 - ▣ Issues with accessibility tools that do not use CSS
-

Images

Definition: Displaying the email address as an image

□ Advantages

- ▣ Spam-bots do not scan images at this time

□ Disadvantages

- ▣ Cannot copy and paste the address
- ▣ User has to retype their address in their mail program, no clickable link
- ▣ Accessibility problems with screen readers and other accessibility tools

Contact Form

Definition: Display a contact form with server-side code responsible for sending email to recipient

□ Advantages

- Email address is never displayed to client
- Email address of recipient is never exposed
- Adding a CAPTCHA key keeps the form secure from auto-populating bots
- Cannot send files or attachments

□ Disadvantages

- User cannot use his preferred e-mail client to compose the message
 - Cannot send files or attachments
-

Other Best Practices

□ Avoid Stereotypes

- Avoid common emails such as: info@domain.com, service@domain.com & admin@domain.com

□ Do Not Respond

- Responding to spammers confirms a valid email

□ Disable HTML

- This expose the user to offensive images
- Can contain web bugs which allow spammers to see that the e-mail address is valid

□ Disposable E-mail addresses

- A temporary address which forwards email to a real account, which the user can disable or abandon
- Use when giving out email to websites